

Sequence Alignments

Q68HV1_MOUSE
 ID Q68HV1_MOUSE PRELIMINARY; PRT; 1925 AA.
 AC Q68HV1;
 DT 11-OCT-2004, integrated into UniProtKB/TrEMBL.
 DT 11-OCT-2004, sequence version 1.
 DT 13-JUN-2006, entry version 11.
 DE Plexin D1.
 GN Name=Plxnd1;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muroidea; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=BALB/c;
 RA Duke-Cohan J.S., Ahmed W., Reinherz E.L.;
 RT "Identification of mouse plexin D1.";
 RL Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC

DR EMBL; AY688678; AAT99561.1; -; mRNA.
 DR UniGene; Mm.3085; -.
 DR Ensembl; ENSMUSG00000030123; Mus musculus.
 DR MGI; MGI:2154244; Plxnd1.
 DR RZPD-ProtExp; IOM17766; -.
 DR GO; GO:0005886; C:plasma membrane; IC.
 DR GO; GO:0005515; F:protein binding; IPI.
 DR GO; GO:0017154; F:semaphorin receptor activity; IPI.
 DR GO; GO:0001569; P:patterning of blood vessels; IMP.
 DR InterPro; IPR002909; IPT_TIG_rcpt.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR013548; Plexin_cytopl.
 DR InterPro; IPR002165; Plexin_repeat.
 DR InterPro; IPR008936; Rho_GAP.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF08337; Plexin_cytopl; 1.
 DR Pfam; PF01437; PSI; 2.
 DR Pfam; PF01403; Sema; 1.
 DR Pfam; PF01833; TIG; 3.
 DR SMART; SM00429; IPT; 3.
 DR SMART; SM00423; PSI; 3.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS51004; SEMA; 1.
 SQ SEQUENCE 1925 AA; 211638 MW; 53E71ADBD05A78FC CRC64;

Query Match 100.0%; Score 9181; DB 2; Length 1925;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db ||||||| 180 SMLNVAANHPNASTVGLVLPPTSGGSRLLVGATYTGFSAFFPRNRSL EDHRFENTPE 239

Qy 61 IAIRSLDARGDLAKLFTFDLNPSDDNILKIKQGAKEQHKLGFVRAFLHPAVPPHSAQPYA 120
||| 240 IAIRSLDARGDLAKLFTFDLNPSDDNILKIKQGAKEQHKLGFVRAFLHPAVPPHSAQPYA 299

Qy 121 YLALNSEARAGDKDSQARSLLARICLPRGAGGDACKLTESYIQLGLQCAGGAGR GDLYSR 180
||| 300 YLALNSEARAGDKDSQARSLLARICLPRGAGGDACKLTESYIQLGLQCAGGAGR GDLYSR 359

Qy 181 LVSVF PAREQFFAVFERPQGAPGARNAPAALCAF RFD DVQAA IRAARTACFVEPAPDVVA 240
||| 360 LVSVF PAREQFFAVFERPQGAPGARNAPAALCAF RFD DVQAA IRAARTACFVEPAPDVVA 419

Qy 241 VLDSVVQGTGPACESKRNIQLQPEQLDCGA AHLQHPLT ILQPLRASP VFRAP GLT AVAVA 300
||| 420 VLDSVVQGTGPACESKRNIQLQPEQLDCGA AHLQHPLT ILQPLRASP VFRAP GLT AVAVA 479

Qy 301 SAN NY TAVFLGTATGR LLKISLNE SMQV VSRRV LTVAY GEPV HVMQ FDPMDPGY LYLM T 360
||| 480 SAN NY TAVFLGTATGR LLKISLNE SMQV VSRRV LTVAY GEPV HVMQ FDPMDPGY LYLM T 539

Qy 361 SHQM ARVKVAACEV HSTCGDCVGAADAYCGWCTLETRCTLQ QDCTNSSQPHFWTSASEGP 420
||| 540 SHQM ARVKVAACEV HSTCGDCVGAADAYCGWCTLETRCTLQ QDCTNSSQPHFWTSASEGP 599

Qy 421 SRC PAMT VLPSEIDVHRDYTG MILQISGSLPSLSGMEMACDYGN GVRTVARVPGP AYDHQ 480
||| 600 SRC PAMT VLPSEIDVHRDYTG MILQISGSLPSLSGMEMACDYGN GVRTVARVPGP AYDHQ 659

Qy 481 IAYCNLLPRAQFPSFPAGQDHVT VEMSVRVKGHNIVS ANFTIYDCSRIGQVYPHTACTSC 540
||| 660 IAYCNLLPRAQFPSFPAGQDHVT VEMSVRVKGHNIVS ANFTIYDCSRIGQVYPHTACTSC 719

Qy 541 LSTQWPCSWCIQLHSCVS NQSQCQDSPNPTSPQDCPQI LPSPLAPVPTGGSQDILVPLTK 600
||| 720 LSTQWPCSWCIQLHSCVS NQSQCQDSPNPTSPQDCPQI LPSPLAPVPTGGSQDILVPLTK 779

Qy 601 ATFFHGSSLECSFGLEESFEAVWANNSLVRCNQVVLHTTQKSQVFPLSLKLKGPPDRFLD 660
||| 780 ATFFHGSSLECSFGLEESFEAVWANNSLVRCNQVVLHTTQKSQVFPLSLKLKGPPDRFLD 839

Qy 661 SPNPMTV VVYNCAMGSPDCS QCLGRE DLGHLCV WNDG CRLRGPLQPLPGT C PAPEIR AIE 720
||| 840 SPNPMTV VVYNCAMGSPDCS QCLGRE DLGHLCV WNDG CRLRGPLQPLPGT C PAPEIR AIE 899

Qy 721 PL SGPLDG GTLLTIRGRNLG RRLSDVAHGVWIG SVACEPLADRTVSEEIVCAT GPAAGA 780
||| 900 PL SGPLDG GTLLTIRGRNLG RRLSDVAHGVWIG SVACEPLADRTVSEEIVCAT GPAAGA 959

Qy 781 FSDVVTVNVSKEGRSREQFSYVLPTVHSLEPSMGP KAGGTRITIH GSDLNVGSM LQVL VN 840
||| 960 FSDVVTVNVSKEGRSREQFSYVLPTVHSLEPSMGP KAGGTRITIH GSDLNVGSM LQVL VN 1019

Qy 841 DTD PCTDLTR TATSITCTVPG TLPS PVCVRF ESRGC VHG NLTFW YM QNPVITA IS PG 900
|||

Db 1020 DTDPCDLTRATTSITCTVPGTLPSPVPCVRFESRGCVHGNLTWYMQNPVITAISPG 1079
Qy 901 RSPVSGGRTITVAGERFHMVQNVSMAVHHIGREPTFCKVLNSTLITCPSPGALSNASAPV 960
|||
Db 1080 RSPVSGGRTITVAGERFHMVQNVSMAVHHIGREPTFCKVLNSTLITCPSPGALSNASAPV 1139
Qy 961 DFFINGRAYADEAAEELLDPAEAQRGSFRLDYLPNPQFSTAKREKWIKHHPGEPLTLVI 1020
|||
Db 1140 DFFINGRAYADEAAEELLDPAEAQRGSFRLDYLPNPQFSTAKREKWIKHHPGEPLTLVI 1199
Qy 1021 HKEQDSLGLSHEYHIKIGQVSCDIQIISDRVIGHCSVNESLGTAEGQLPITIQVGNFNQT 1080
|||
Db 1200 HKEQDSLGLSHEYHIKIGQVSCDIQIISDRVIGHCSVNESLGTAEGQLPITIQVGNFNQT 1259
Qy 1081 IATLQLGGSETAIVVSIVICSVLLLLSVVALFVFCTKSRRAYERWQKTLQMEEMESQIR 1140
|||
Db 1260 IATLQLGGSETAIVVSIVICSVLLLLSVVALFVFCTKSRRAYERWQKTLQMEEMESQIR 1319
Qy 1141 EEIRKGFAELQTDMDTDLTKELNRSGQIPFLEYKHFVTRTFPKCSSLYEERYVLPSKTLN 1200
|||
Db 1320 EEIRKGFAELQTDMDTDLTKELNRSGQIPFLEYKHFVTRTFPKCSSLYEERYVLPSKTLN 1379
Qy 1201 SQGGSPPQETHPLLGEWNIPHCPRSMEEGISLFSSLLNNKHFLIVFVHALEQQKDFAVR 1260
|||
Db 1380 SQGGSPPQETHPLLGEWNIPHCPRSMEEGISLFSSLLNNKHFLIVFVHALEQQKDFAVR 1439
Qy 1261 DRCSLASLLTIALHGKLEYYTSIMKELLVDLIDASAACKPKLMLRRTESVVEKMLTNWMS 1320
|||
Db 1440 DRCSLASLLTIALHGKLEYYTSIMKELLVDLIDASAACKPKLMLRRTESVVEKMLTNWMS 1499
Qy 1321 ICMYGCLRETVGEPFFLLLCAIKQQINKGSIDAITGKARYTLNEEWLLRENIEAKPRNLN 1380
|||
Db 1500 ICMYGCLRETVGEPFFLLLCAIKQQINKGSIDAITGKARYTLNEEWLLRENIEAKPRNLN 1559
Qy 1381 VSFQCGMDSLSVRAMDTDTLTQVKKEKILEAFCKNVPYSQWPRAEDVDLEWFASSTQSYV 1440
|||
Db 1560 VSFQCGMDSLSVRAMDTDTLTQVKKEKILEAFCKNVPYSQWPRAEDVDLEWFASSTQSYV 1619
Qy 1441 LRDLDDTSVVEDGRKKLNTLAHYKIPEGASLAMS LTDKDSTLGRVKLDTEKYFHLVLP 1500
|||
Db 1620 LRDLDDTSVVEDGRKKLNTLAHYKIPEGASLAMS LTDKDSTLGRVKLDTEKYFHLVLP 1679
Qy 1501 TDELVEPKKSHRQSHRKVLP E IYLTRL LSTKGT LQKF LDDLF KAIL SIRE DKPL AVKY 1560
|||
Db 1680 TDELVEPKKSHRQSHRKVLP E IYLTRL LSTKGT LQKF LDDLF KAIL SIRE DKPL AVKY 1739
Qy 1561 FFDLLEEQAERKGI SD PDL HI WKT NSLPL RFVN ILKNP QFV FD IE KTD HIDA CLS VIA 1620
|||
Db 1740 FFDLLEEQAERKGI SD PDL HI WKT NSLPL RFVN ILKNP QFV FD IE KTD HIDA CLS VIA 1799
Qy 1621 QAFIDACSI S DLQ LGK DS PTN KLL YAKEI PEY RKT VQ RYY KQI QDM TPL SE QEM NAH LAE 1680
|||
Db 1800 QAFIDACSI S DLQ LGK DS PTN KLL YAKEI PEY RKT VQ RYY KQI QDM TPL SE QEM NAH LAE 1859
Qy 1681 ESRKYQNEFNTNVAMA E IY KYAK RYRP QIMA ALE ANPTARRTQLQHKFEQVVALMEN NIY 1740
|||
Db 1860 ESRKYQNEFNTNVAMA E IY KYAK RYRP QIMA ALE ANPTARRTQLQHKFEQVVALMEN NIY 1919

Qy 1741 ECYSEA 1746
| | | | |
Db 1920 ECYSEA 1925

RESULT 4

PLXD1_HUMAN

ID PLXD1_HUMAN STANDARD; PRT; 1925 AA.

AC Q9Y4D7; Q6PJS9; Q8IZJ2; Q9BTQ2;

DT 31-AUG-2004, integrated into UniProtKB/Swiss-Prot.

DT 31-AUG-2004, sequence version 2.

DT 25-JUL-2006, entry version 36.

DE Plexin-D1 precursor.

GN Name=PLXND1; Synonyms=KIAA0620;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;

OC Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE mRNA] (ISOFORM 1).

RC TISSUE=Brain;

RX MEDLINE=98403880; PubMed=9734811; DOI=10.1093/dnare/5.3.169;

RA Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes. X. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.";

RL DNA Res. 5:169-176(1998).

RN [2]

RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), AND TISSUE SPECIFICITY.

RX MEDLINE=22299888; PubMed=12412018; DOI=10.1002/dvdy.10159;

RA van der Zwaag B., Hellmons A.J.C.G.M., Leenders W.P.J., Burbach J.P.H., Brunner H.G., Padberg G.W., Van Bokhoven H.;

RT "PLEXIN-D1, a novel plexin family member, is expressed in vascular endothelium and the central nervous system during mouse embryogenesis.";

RL Dev. Dyn. 225:336-343(2002).

RN [3]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE mRNA] OF 1386-1925 (ISOFORMS 1 AND 2).

RC TISSUE=Muscle, and Uterus;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP GLYCOSYLATION [LARGE SCALE ANALYSIS] AT ASN-500, AND MASS
RP SPECTROMETRY.
RX PubMed=16335952; DOI=10.1021/pr0502065;
RA Liu T., Qian W.-J., Gritsenko M.A., Camp D.G. II, Monroe M.E.,
RA Moore R.J., Smith R.D.;
RT "Human plasma N-glycoproteome analysis by immunoaffinity subtraction,
RT hydrazide chemistry, and mass spectrometry.";
RL J. Proteome Res. 4:2070-2080(2005).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9Y4D7-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9Y4D7-2; Sequence=VSP_011516;
CC -!- TISSUE SPECIFICITY: Detected at low levels in heart, placenta,
CC lung, skeletal muscle, kidney, thymus and liver. Detected at very
CC low levels in brain, colon, spleen, small intestine and peripheral
CC blood leukocytes.
CC -!- SIMILARITY: Belongs to the plexin family.
CC -!- SIMILARITY: Contains 3 IPT/TIG domains.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AB014520; BAA31595.1; ALT_INIT; mRNA.
DR EMBL; AY116661; AAM49063.1; -; mRNA.
DR EMBL; BC003526; AAH03526.1; -; mRNA.
DR EMBL; BC011848; AAH11848.1; -; mRNA.
DR UniGene; Hs.301685; -.
DR Ensembl; ENSG00000004399; Homo sapiens.
DR HGNC; HGNC:9107; PLXND1.
DR MIM; 604282; gene.
DR InterPro; IPR002909; IPT_TIG_rcpt.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR013548; Plexin_cytopl.
DR InterPro; IPR002165; Plexin_repeat.
DR InterPro; IPR008936; Rho_GAP.
DR InterPro; IPR001627; Sema.
DR Pfam; PF08337; Plexin_cytopl; 1.
DR Pfam; PF01437; PSI; 2.
DR Pfam; PF01403; Sema; 1.
DR Pfam; PF01833; TIG; 3.
DR SMART; SM00429; IPT; 3.
DR SMART; SM00423; PSI; 3.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS51004; SEMA; 1.
KW Alternative splicing; Glycoprotein; Membrane; Polymorphism; Receptor;
KW Repeat; Signal; Transmembrane.
FT SIGNAL 1 46 Potential.
FT CHAIN 47 1925 Plexin-D1.
FT /FTId=PRO_0000024676.
FT TOPO_DOM 47 1271 Extracellular (Potential).
FT TRANSMEM 1272 1292 Potential.
FT TOPO_DOM 1293 1925 Cytoplasmic (Potential).

FT	DOMAIN	47	546	Sema.
FT	DOMAIN	891	979	IPT/TIG 1.
FT	DOMAIN	981	1066	IPT/TIG 2.
FT	DOMAIN	1069	1160	IPT/TIG 3.
FT	CARBOHYD	86	86	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	155	155	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	188	188	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	224	224	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	481	481	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	500	500	N-linked (GlcNAc. . .).
FT	CARBOHYD	583	583	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	696	696	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	736	736	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	802	802	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	965	965	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1017	1017	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1060	1060	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1099	1099	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1118	1118	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1132	1132	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1237	1237	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1257	1257	N-linked (GlcNAc. . .) (Potential).
FT	VAR_SEQ	1766	1925	SLPLRFWVNILKNPQFVFDIDKTDHIDACLSVIAQAFIDAC SISDLQLGKDSPTNKLLYAKEIPEYRKIVQRYYKQIQDMTP LSEQEMNAHLAEESRKYQNEFNTNVAMAEIYKYAKRYRPQI MAALEANPTARRTQLQHKFEQVVALMEDNIYECYSEA -> RWRPSSPVVLGEHPEPPVCL (in isoform 2). /FTId=VSP_011516.
FT	VARIANT	870	870	M -> V (in dbSNP:2255703). /FTId=VAR_022144.
SQ	SEQUENCE	1925 AA;	212095 MW;	26001F5D0B2A80E5 CRC64;

Qy	1	SMLNVAANHPNASTVGLVLPPSGTGGSRLLVGATYTGFSAFFPRNRSLEDHRFENTPE	60
		: : : : : : : : :	
Db	178	SMLNVAANHPNASTVGLVLPPAAGAGGSRLLVGATYTGYGSSFFPRNRSLEDHRFENTPE	237
Qy	61	IAIRSLDARGDLAKLFTFDLNPSDDNILKIKQGAKEQHKLGFVRAFLHPAVPPHSAQPYA	120
		: : : : : : : : :	
Db	238	IAIRSLDTRGDLAKLFTFDLNPSDDNILKIKQGAKEQHKLGFVSAFLHPSDPPPQGAQSYA	297
Qy	121	YLALNSEARAGDKDSQARSLLARICLPRGAGGDAKKLTESYIQLGLQCAGGAGRGDLYSR	180
		: : : : : : : : : : :	
Db	298	YLALNSEARAGDKESQARSLLARICLPHGAGGDAKKLTESYIQLGLQCAGGAGRGDLYSR	357
Qy	181	LVSVF PAREQFFAVFERPQGAPGARNAPAALCAF RF DDVQAAIRAARTACFVEPAPDVVA	240
		: : : : : : : : : : :	
Db	358	LVSVF PARERLF AVFERPQGSPAARAAPAALCAF RF ADVRAAIRAARTACFVEPAPDVVA	417
Qy	241	VLD SVV QGTGPACESKRNIQLQPEQLDCGA AHLQHPLT ILQPLRASP VFRAPGLTAVAVA	300
		: : : : : : : : : :	
Db	418	VLD SVV QGTGPACERKLN IQLQPEQLDCGA AHLQHPLS ILQPLKATPVFRAPGLT SVAVA	477

Qy 301 SANNYTAVFLGTATGRLKISLNESMQVVSRRVLTVAYGEPVHHVMQFDPMDPGYLYLMT 360
| ||||||| | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 478 SVNNYTAVFLGTVNGRLLKINLNESMQVVSRRVTVAYGEPVHHVMQFDPADSVLYLMT 537

Qy 361 SHQMARVKVAACEVHSTCGDCVGAADAYCGWCTLETRCTLQQDCTNSSQPHFWTSASEGP 420
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 538 SHQMARVKVAACNVHSTCGDCVGAADAYCGWCALETRCTLQQDCTNSSQPHFWTSASEGP 597

Qy 421 SRCPAMTVPSEIDVHRDVTGMILQISGSLPSLSGMEMACDYGNVRTVARVPGPAYDHQ 480
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 598 SRCPAMTVPSEIDVRQEYPMILQISGSLPSLSGMEMACDYGNNIRTVARVPGPAFGHQ 657

Qy 481 IAYCNLLPRAQFPSFPAGQDHVTVEMSVRVKGHNIVSANFTIYDCSRIGQVYPHTACTSC 540
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 658 IAYCNLLPRDQFPPFPQNQDHVTVEMSVRVNGRNIVKANFTIYDCSRTAQVYVPHACTSC 717

Qy 541 LSTQWPCSWC1QLHSCVSNQSQCQDSPNPTSPQDCPQILPSPLAPVPTGGSQDILVPLTK 600
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 718 LSAQWPCFWCSQQHSCVSNQSRCEASPNPTSPQDCPRTLLSPLAPVPTGGSQNILVPLAN 777

Qy 601 ATFFHGSSLECSFGLEESFEAWWANNSLVRNCQVVLHTTQKSQVFPLSLKLKGPPDRFLD 660
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 778 TAFFQGAALECSFGLEEIFEAVWVNESVVRCDQVVLHTTRKSQVFPLSQLKGRPARFLD 837

Qy 661 SPNPMTVVVYNCAMGSPDCSQCLGREDLGHLCVWNDGCRRLGPLQPLPGTCPAPEIRAE 720
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 838 SPEPMTVMVYNCAMGSPDCSQCLGREDLGHLCMWDGCRRLGPLQPMAGTCPAPEIRAE 897

Qy 721 PLSGPLDGTTLLTIRGRNLGRRLSDAHGVWIGSVACEPLADRYTVSEEIVCATGPAAGA 780
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 898 PLSGPLDGTTLLTIRGRNLGRRLSDAHGVWIGGVACEPLPDRTVSEEIVCVTGPAPGP 957

Qy 781 FSDVVTVNVSKEGRSREQFSYVLPTVHSLEPSMGPKAGGTRITIHGSDLNVGSMQLQVLVN 840
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 958 LSGVVTVNASKEGKSRDRFSYVLPVHSLEPTMGPKAGGTRITIHGNDLHVGSSELQVLVN 1017

Qy 841 DTDPCDLTRTATSITCTVPGGTLPSPVPVCVRFESRGCVHGNLTWYMQNPVITAISPG 900
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1018 DTDPCTELRTDTSIACTMPEGALPAPPVPCVRFERRGCVHGNLTWYMQNPVITAISPR 1077

Qy 901 RSPVSGGRTITVAGERFHMVQNVSMAVHHIGREPTFKVLNSTLITCPSPGALSNASAPV 960
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1078 RSPVSGGRTITVAGERFHMVQNVSMAVHHIGRETLCKVLNSTLITCPSPGALSNASAPV 1137

Qy 961 DFFINGRAYADE--AAEELLDPAEAQRGSRFRLDYLPNPQFSTAKREKWIKHPGEPLTL 1018
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1138 DFFINGRAYADEVAVAEEELLDPEEAQRGSRFRLDYLPNPQFSTAKREKWIKHPGEPLTL 1197

Qy 1019 VIHKEQDSLGLSESHEYHIKIGQVSCDIQIIISDRVIHCSVNESLGTAEGQLPITIQVGNFN 1078
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1198 VIHKEQDSLGLQSHEYRVKIGQVSCDIQIVSDRIIHCSVNESLGAAVGQLPITIQVGNFN 1257

Qy 1079 QTIATLQLGGSETAIVVSIVICSVLLLLSVVALFVFCTKSRAERYWQKTLQMEEMESQ 1138
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1258 QTIATLQLGGSETAIIIVSIVICSVLLLLSVVALFVFCTKSRAERYWQKTLQMEEMESQ 1317

Qy 1139 IREEIRKGFAELQTDMDLTKELENRSQGIPFLEYKHFVTRTFPKCSSLYEERYVLPSKT 1198
||| ||| ||| ||| ||| ||| ||| :|
Db 1318 IREEIRKGFAELQTDMDLTKELENRSQGIPFLEYKHFVTRTFPKCSSLYEERYVLPSQT 1377

Qy 1199 LNSQGGSPPQETHPLLGEWNIPEHCRPSMEEGISLFSSLLNNKHFLIVFVHALEQQKDFA 1258
||| ||| | ||| ||| ||| :||| ||| ||| ||| |||
Db 1378 LNSQGSSQAQETHPLLGEWKIPESCRPNMEEGISLFSSLLNNKHFLIVFVHALEQQKDFA 1437

Qy 1259 VRDRCSLASLLTIALHGKLEYYTSIMKELLVDLIDASAAKNPKLMRRTESVVEKMLTNW 1318
||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1438 VRDRCSLASLLTIALHGKLEYYTSIMKELLVDLIDASAAKNPKLMRRTESVVEKMLTNW 1497

Qy 1319 MSICMYGCLRETVGEPFLLLCAIKQQINKGSIDAITGKARYTLNEEWLLRENIEAKPRN 1378
||| ||| ||| ||| ||| ||| ||| |||
Db 1498 MSICMYSCLRETVGEPFLLLCAIKQQINKGSIDAITGKARYTLNEEWLLRENIEAKPRN 1557

Qy 1379 LNVSFQGCGMDSLSVRAMDTDTLTQVKEKILEAFCKNVPYSQWPRAEDVDLEWFASSTQS 1438
||| ||| ||| ||| ||| ||| |||
Db 1558 LNVSFQGCGMDSLSVRAMDTDTLTQVKEKILEAFCKNVPYSQWPRAEDVDLEWFASSTQS 1617

Qy 1439 YVLRDLDTSVVEDGRKKLNTLAHYKIPEGASLAMS LTD KKD STLGRVKDLDEKYFHLV 1498
|:||| ||| ||| ||| ||| ||| ||| ||| ||| :||| ||| |||
Db 1618 YILRDLDTSVVEDGRKKLNTLAHYKIPEGASLAMS ID KKD NTLGRVKDLDEKYFHLV 1677

Qy 1499 LPTDELVEPKKSHRQSHRKVLPEIYLTRLSTKGLQKF LDDLFKA IL SIRE DKPPLAV 1558
||| ||| ||| ||| ||| ||| |||
Db 1678 LPTDELAEPKKSHRQSHRKVLPEIYLTRLSTKGLQKF LDDLFKA IL SIRE DKPPLAV 1737

Qy 1559 KYFFDFLEEQAEKRGISDPDTLHIWKTNSLPLRFWVNILKNPQFVFDIEKTDHIDACLSV 1618
||| ||| ||| ||| ||| ||| ||| ||| :||| |||
Db 1738 KYFFDFLEEQAEKRGISDPDTLHIWKTNSLPLRFWVNILKNPQFVFDIDKTDHIDACLSV 1797

Qy 1619 IAQAFIDACSISDLQLGKDSPTNKLLYAKEIPEYRKT VQRYYKQIQDMTPLSEQEMNAHL 1678
||| ||| ||| ||| ||| ||| |||
Db 1798 IAQAFIDACSISDLQLGKDSPTNKLLYAKEIPEYRKT VQRYYKQIQDMTPLSEQEMNAHL 1857

Qy 1679 AEESRKYQNEFTNVAMAEIYKYAKRYRPQIMAALLEANPTARRTQLQHKFEQVVALMENN 1738
||| ||| ||| ||| ||| ||| ||| :|
Db 1858 AEESRKYQNEFTNVAMAEIYKYAKRYRPQIMAALLEANPTARRTQLQHKFEQVVALMEDN 1917

Qy 1739 IYECYSEA 1746
||| |||
Db 1918 IYECYSEA 1925